

UAB THE HACK! 2025 - WiFi Dataset Analysis

Nivel 1: ROOKIE - Análisis Básico

Objetivo: Explorar y visualizar el dataset de WiFi de la UAB

Dataset:

- Access Points (APs): 2,333 archivos con snapshots de APs del campus
 - Clientes: 3,199 archivos con dispositivos conectados
 - Período: Abril-Julio 2025
-



1. Importar Librerías

Primero, importamos las librerías necesarias para el análisis.

```
In [3]: # Librerías estándar y científicas
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from pathlib import Path
import warnings

# Configuración general
warnings.filterwarnings("ignore")
sns.set_theme(style="whitegrid", palette="pastel")
plt.rcParams.update({
    "figure.figsize": (12, 6),
    "axes.titlesize": 18,
    "axes.labelsize": 14
})

# Carga de funciones personalizadas
from utils.data_loader import (
    load_aps,
    load_clients,
    print_dataset_summary,
    get_top_aps
)

print("✅ Librerías y utilidades cargadas correctamente")
```

✅ Librerías y utilidades cargadas correctamente




2. Cargar Datos

Cargamos una muestra del dataset. Empezaremos con 10 archivos para exploración rápida.

Nota: Para cargar TODO el dataset, cambia `max_files=None`

```
In [4]: # Cargar Access Points (primeros 10 archivos)
df_aps = load_aps(
    data_dir="../../anonymized_data/aps",
    max_files=None,
    verbose=True
)

print("\n" + "="*60)
print(f"🌀 Aps cargados: {len(df_aps):,} registros")
print("="*60)
```

 Encontrados 2333 archivos en ..\anonymized_data\aps
 Cargando todos archivos...

Procesados	10/2333	archivos...	(11280 registros)
Procesados	20/2333	archivos...	(22560 registros)
Procesados	30/2333	archivos...	(33834 registros)
Procesados	40/2333	archivos...	(45119 registros)
Procesados	50/2333	archivos...	(56409 registros)
Procesados	60/2333	archivos...	(67699 registros)
Procesados	70/2333	archivos...	(79003 registros)
Procesados	80/2333	archivos...	(90319 registros)
Procesados	90/2333	archivos...	(101630 registros)
Procesados	100/2333	archivos...	(112940 registros)
Procesados	110/2333	archivos...	(124250 registros)
Procesados	120/2333	archivos...	(135593 registros)
Procesados	130/2333	archivos...	(147013 registros)
Procesados	140/2333	archivos...	(158433 registros)
Procesados	150/2333	archivos...	(169853 registros)
Procesados	160/2333	archivos...	(181273 registros)
Procesados	170/2333	archivos...	(192693 registros)
Procesados	180/2333	archivos...	(204113 registros)
Procesados	190/2333	archivos...	(215533 registros)
Procesados	200/2333	archivos...	(226953 registros)
Procesados	210/2333	archivos...	(238373 registros)
Procesados	220/2333	archivos...	(249793 registros)
Procesados	230/2333	archivos...	(261213 registros)
Procesados	240/2333	archivos...	(272633 registros)
Procesados	250/2333	archivos...	(284053 registros)
Procesados	260/2333	archivos...	(295473 registros)
Procesados	270/2333	archivos...	(306893 registros)
Procesados	280/2333	archivos...	(318313 registros)
Procesados	290/2333	archivos...	(329733 registros)
Procesados	300/2333	archivos...	(341153 registros)
Procesados	310/2333	archivos...	(352573 registros)
Procesados	320/2333	archivos...	(363993 registros)
Procesados	330/2333	archivos...	(375413 registros)
Procesados	340/2333	archivos...	(386833 registros)
Procesados	350/2333	archivos...	(398253 registros)
Procesados	360/2333	archivos...	(409673 registros)
Procesados	370/2333	archivos...	(421093 registros)
Procesados	380/2333	archivos...	(432513 registros)
Procesados	390/2333	archivos...	(443933 registros)
Procesados	400/2333	archivos...	(455353 registros)
Procesados	410/2333	archivos...	(466773 registros)
Procesados	420/2333	archivos...	(478193 registros)
Procesados	430/2333	archivos...	(489613 registros)
Procesados	440/2333	archivos...	(501033 registros)
Procesados	450/2333	archivos...	(512453 registros)
Procesados	460/2333	archivos...	(523873 registros)
Procesados	470/2333	archivos...	(535293 registros)
Procesados	480/2333	archivos...	(546713 registros)
Procesados	490/2333	archivos...	(558133 registros)
Procesados	500/2333	archivos...	(569553 registros)
Procesados	510/2333	archivos...	(580973 registros)
Procesados	520/2333	archivos...	(592393 registros)
Procesados	530/2333	archivos...	(603813 registros)
Procesados	540/2333	archivos...	(615233 registros)
Procesados	550/2333	archivos...	(626653 registros)
Procesados	560/2333	archivos...	(638073 registros)
Procesados	570/2333	archivos...	(649493 registros)
Procesados	580/2333	archivos...	(660913 registros)

Procesados 590/2333 archivos... (672333 registros)
Procesados 600/2333 archivos... (683753 registros)
Procesados 610/2333 archivos... (695173 registros)
Procesados 620/2333 archivos... (706593 registros)
Procesados 630/2333 archivos... (718013 registros)
Procesados 640/2333 archivos... (729433 registros)
Procesados 650/2333 archivos... (740853 registros)
Procesados 660/2333 archivos... (752273 registros)
Procesados 670/2333 archivos... (763693 registros)
Procesados 680/2333 archivos... (775113 registros)
Procesados 690/2333 archivos... (786533 registros)
Procesados 700/2333 archivos... (797953 registros)
Procesados 710/2333 archivos... (809373 registros)
Procesados 720/2333 archivos... (820793 registros)
Procesados 730/2333 archivos... (832213 registros)
Procesados 740/2333 archivos... (843633 registros)
Procesados 750/2333 archivos... (855053 registros)
Procesados 760/2333 archivos... (866473 registros)
Procesados 770/2333 archivos... (877893 registros)
Procesados 780/2333 archivos... (889313 registros)
Procesados 790/2333 archivos... (900733 registros)
Procesados 800/2333 archivos... (912153 registros)
Procesados 810/2333 archivos... (923573 registros)
Procesados 820/2333 archivos... (934993 registros)
Procesados 830/2333 archivos... (946413 registros)
Procesados 840/2333 archivos... (957833 registros)
Procesados 850/2333 archivos... (969253 registros)
Procesados 860/2333 archivos... (980673 registros)
Procesados 870/2333 archivos... (992093 registros)
Procesados 880/2333 archivos... (1003513 registros)
Procesados 890/2333 archivos... (1014933 registros)
Procesados 900/2333 archivos... (1026353 registros)
Procesados 910/2333 archivos... (1037773 registros)
Procesados 920/2333 archivos... (1049193 registros)
Procesados 930/2333 archivos... (1060613 registros)
Procesados 940/2333 archivos... (1072033 registros)
Procesados 950/2333 archivos... (1083453 registros)
Procesados 960/2333 archivos... (1094873 registros)
Procesados 970/2333 archivos... (1106293 registros)
Procesados 980/2333 archivos... (1117713 registros)
Procesados 990/2333 archivos... (1129133 registros)
Procesados 1000/2333 archivos... (1140553 registros)
Procesados 1010/2333 archivos... (1151973 registros)
Procesados 1020/2333 archivos... (1163393 registros)
Procesados 1030/2333 archivos... (1174813 registros)
Procesados 1040/2333 archivos... (1186233 registros)
Procesados 1050/2333 archivos... (1197653 registros)
Procesados 1060/2333 archivos... (1209073 registros)
Procesados 1070/2333 archivos... (1220493 registros)
Procesados 1080/2333 archivos... (1231913 registros)
Procesados 1090/2333 archivos... (1243333 registros)
Procesados 1100/2333 archivos... (1254753 registros)
Procesados 1110/2333 archivos... (1266173 registros)
Procesados 1120/2333 archivos... (1277593 registros)
Procesados 1130/2333 archivos... (1289013 registros)
Procesados 1140/2333 archivos... (1300433 registros)
Procesados 1150/2333 archivos... (1311853 registros)
Procesados 1160/2333 archivos... (1323273 registros)
Procesados 1170/2333 archivos... (1334693 registros)
Procesados 1180/2333 archivos... (1346113 registros)

Procesados	1190/2333	archivos...	(1357533 registros)
Procesados	1200/2333	archivos...	(1368953 registros)
Procesados	1210/2333	archivos...	(1380373 registros)
Procesados	1220/2333	archivos...	(1391793 registros)
Procesados	1230/2333	archivos...	(1403213 registros)
Procesados	1240/2333	archivos...	(1414633 registros)
Procesados	1250/2333	archivos...	(1426053 registros)
Procesados	1260/2333	archivos...	(1437473 registros)
Procesados	1270/2333	archivos...	(1448893 registros)
Procesados	1280/2333	archivos...	(1460313 registros)
Procesados	1290/2333	archivos...	(1471733 registros)
Procesados	1300/2333	archivos...	(1483153 registros)
Procesados	1310/2333	archivos...	(1494573 registros)
Procesados	1320/2333	archivos...	(1505993 registros)
Procesados	1330/2333	archivos...	(1517413 registros)
Procesados	1340/2333	archivos...	(1528833 registros)
Procesados	1350/2333	archivos...	(1540253 registros)
Procesados	1360/2333	archivos...	(1551673 registros)
Procesados	1370/2333	archivos...	(1563093 registros)
Procesados	1380/2333	archivos...	(1574513 registros)
Procesados	1390/2333	archivos...	(1585933 registros)
Procesados	1400/2333	archivos...	(1597353 registros)
Procesados	1410/2333	archivos...	(1608773 registros)
Procesados	1420/2333	archivos...	(1620193 registros)
Procesados	1430/2333	archivos...	(1631613 registros)
Procesados	1440/2333	archivos...	(1643033 registros)
Procesados	1450/2333	archivos...	(1654453 registros)
Procesados	1460/2333	archivos...	(1665873 registros)
Procesados	1470/2333	archivos...	(1677293 registros)
Procesados	1480/2333	archivos...	(1688713 registros)
Procesados	1490/2333	archivos...	(1700133 registros)
Procesados	1500/2333	archivos...	(1711553 registros)
Procesados	1510/2333	archivos...	(1722973 registros)
Procesados	1520/2333	archivos...	(1734393 registros)
Procesados	1530/2333	archivos...	(1745813 registros)
Procesados	1540/2333	archivos...	(1757233 registros)
Procesados	1550/2333	archivos...	(1768653 registros)
Procesados	1560/2333	archivos...	(1780073 registros)
Procesados	1570/2333	archivos...	(1791493 registros)
Procesados	1580/2333	archivos...	(1802913 registros)
Procesados	1590/2333	archivos...	(1814333 registros)
Procesados	1600/2333	archivos...	(1825753 registros)
Procesados	1610/2333	archivos...	(1837173 registros)
Procesados	1620/2333	archivos...	(1848593 registros)
Procesados	1630/2333	archivos...	(1860013 registros)
Procesados	1640/2333	archivos...	(1871433 registros)
Procesados	1650/2333	archivos...	(1882853 registros)
Procesados	1660/2333	archivos...	(1894273 registros)
Procesados	1670/2333	archivos...	(1905693 registros)
Procesados	1680/2333	archivos...	(1917115 registros)
Procesados	1690/2333	archivos...	(1928545 registros)
Procesados	1700/2333	archivos...	(1939975 registros)
Procesados	1710/2333	archivos...	(1951405 registros)
Procesados	1720/2333	archivos...	(1962835 registros)
Procesados	1730/2333	archivos...	(1974265 registros)
Procesados	1740/2333	archivos...	(1985695 registros)
Procesados	1750/2333	archivos...	(1997125 registros)
Procesados	1760/2333	archivos...	(2008555 registros)
Procesados	1770/2333	archivos...	(2019985 registros)
Procesados	1780/2333	archivos...	(2031415 registros)

Procesados 1790/2333 archivos... (2042845 registros)
Procesados 1800/2333 archivos... (2054275 registros)
Procesados 1810/2333 archivos... (2065705 registros)
Procesados 1820/2333 archivos... (2077135 registros)
Procesados 1830/2333 archivos... (2088565 registros)
Procesados 1840/2333 archivos... (2099995 registros)
Procesados 1850/2333 archivos... (2111425 registros)
Procesados 1860/2333 archivos... (2122855 registros)
Procesados 1870/2333 archivos... (2134285 registros)
Procesados 1880/2333 archivos... (2145715 registros)
Procesados 1890/2333 archivos... (2157145 registros)
Procesados 1900/2333 archivos... (2168575 registros)
Procesados 1910/2333 archivos... (2180005 registros)
Procesados 1920/2333 archivos... (2191435 registros)
Procesados 1930/2333 archivos... (2202865 registros)
Procesados 1940/2333 archivos... (2214295 registros)
Procesados 1950/2333 archivos... (2225725 registros)
Procesados 1960/2333 archivos... (2237155 registros)
Procesados 1970/2333 archivos... (2248585 registros)
Procesados 1980/2333 archivos... (2260015 registros)
Procesados 1990/2333 archivos... (2271445 registros)
Procesados 2000/2333 archivos... (2282875 registros)
Procesados 2010/2333 archivos... (2294305 registros)
Procesados 2020/2333 archivos... (2305735 registros)
Procesados 2030/2333 archivos... (2317165 registros)
Procesados 2040/2333 archivos... (2328589 registros)
Procesados 2050/2333 archivos... (2340009 registros)
Procesados 2060/2333 archivos... (2351429 registros)
Procesados 2070/2333 archivos... (2362849 registros)
Procesados 2080/2333 archivos... (2374269 registros)
Procesados 2090/2333 archivos... (2385689 registros)
Procesados 2100/2333 archivos... (2397109 registros)
Procesados 2110/2333 archivos... (2408529 registros)
Procesados 2120/2333 archivos... (2419949 registros)
Procesados 2130/2333 archivos... (2431369 registros)
Procesados 2140/2333 archivos... (2442789 registros)
Procesados 2150/2333 archivos... (2454209 registros)
Procesados 2160/2333 archivos... (2465629 registros)
Procesados 2170/2333 archivos... (2477049 registros)
Procesados 2180/2333 archivos... (2488469 registros)
Procesados 2190/2333 archivos... (2499889 registros)
Procesados 2200/2333 archivos... (2511309 registros)
Procesados 2210/2333 archivos... (2522729 registros)
Procesados 2220/2333 archivos... (2534149 registros)
Procesados 2230/2333 archivos... (2545569 registros)
Procesados 2240/2333 archivos... (2556989 registros)
Procesados 2250/2333 archivos... (2568409 registros)
Procesados 2260/2333 archivos... (2579829 registros)
Procesados 2270/2333 archivos... (2591249 registros)
Procesados 2280/2333 archivos... (2602669 registros)
Procesados 2290/2333 archivos... (2614089 registros)
Procesados 2300/2333 archivos... (2625509 registros)
Procesados 2310/2333 archivos... (2636929 registros)
Procesados 2320/2333 archivos... (2648349 registros)
Procesados 2330/2333 archivos... (2659769 registros)



Cargados 2663195 registros de 2333 archivos



Memoria: 3552.10 MB



=====

🌀 APs cargados: 2,663,195 registros

=====

```
In [5]: # Cargar Clientes (primeros 10 archivos)
df_clients = load_clients(
    data_dir="../../anonymized_data/clients",
    max_files=None,
    verbose=True
)

print("\n" + "="*60)
print(f"🌀 Clientes cargados: {len(df_clients):,} registros")
print("="*60)
```

 Encontrados 3199 archivos en ..\anonymized_data\clients
 Cargando todos archivos...

Procesados	10/3199	archivos...	(11519 registros)
Procesados	20/3199	archivos...	(116418 registros)
Procesados	30/3199	archivos...	(250458 registros)
Procesados	40/3199	archivos...	(328784 registros)
Procesados	50/3199	archivos...	(337106 registros)
Procesados	60/3199	archivos...	(411981 registros)
Procesados	70/3199	archivos...	(528889 registros)
Procesados	80/3199	archivos...	(599936 registros)
Procesados	90/3199	archivos...	(610104 registros)
Procesados	100/3199	archivos...	(622012 registros)
Procesados	110/3199	archivos...	(632322 registros)
Procesados	120/3199	archivos...	(640318 registros)
Procesados	130/3199	archivos...	(651654 registros)
Procesados	140/3199	archivos...	(659062 registros)
Procesados	150/3199	archivos...	(736726 registros)
Procesados	160/3199	archivos...	(868644 registros)
Procesados	170/3199	archivos...	(960781 registros)
Procesados	180/3199	archivos...	(973164 registros)
Procesados	190/3199	archivos...	(1034007 registros)
Procesados	200/3199	archivos...	(1173902 registros)
Procesados	210/3199	archivos...	(1280601 registros)
Procesados	220/3199	archivos...	(1303127 registros)
Procesados	230/3199	archivos...	(1338001 registros)
Procesados	240/3199	archivos...	(1446465 registros)
Procesados	250/3199	archivos...	(1523066 registros)
Procesados	260/3199	archivos...	(1556703 registros)
Procesados	270/3199	archivos...	(1578912 registros)
Procesados	280/3199	archivos...	(1695006 registros)
Procesados	290/3199	archivos...	(1814744 registros)
Procesados	300/3199	archivos...	(1866190 registros)
Procesados	310/3199	archivos...	(1877513 registros)
Procesados	320/3199	archivos...	(1958587 registros)
Procesados	330/3199	archivos...	(2053506 registros)
Procesados	340/3199	archivos...	(2099387 registros)
Procesados	350/3199	archivos...	(2107511 registros)
Procesados	360/3199	archivos...	(2116598 registros)
Procesados	370/3199	archivos...	(2124828 registros)
Procesados	380/3199	archivos...	(2132954 registros)
Procesados	390/3199	archivos...	(2141133 registros)
Procesados	400/3199	archivos...	(2148918 registros)
Procesados	410/3199	archivos...	(2159753 registros)
Procesados	420/3199	archivos...	(2172557 registros)
Procesados	430/3199	archivos...	(2183416 registros)
Procesados	440/3199	archivos...	(2189722 registros)
Procesados	450/3199	archivos...	(2189742 registros)
Procesados	460/3199	archivos...	(2189762 registros)
Procesados	470/3199	archivos...	(2190019 registros)
Procesados	480/3199	archivos...	(2191818 registros)
Procesados	490/3199	archivos...	(2196212 registros)
Procesados	500/3199	archivos...	(2203166 registros)
Procesados	510/3199	archivos...	(2210198 registros)
Procesados	520/3199	archivos...	(2214133 registros)
Procesados	530/3199	archivos...	(2218669 registros)
Procesados	540/3199	archivos...	(2223628 registros)
Procesados	550/3199	archivos...	(2226646 registros)
Procesados	560/3199	archivos...	(2230904 registros)
Procesados	570/3199	archivos...	(2233982 registros)
Procesados	580/3199	archivos...	(2237539 registros)

Procesados	590/3199	archivos...	(2241180 registros)
Procesados	600/3199	archivos...	(2244260 registros)
Procesados	610/3199	archivos...	(2248230 registros)
Procesados	620/3199	archivos...	(2251178 registros)
Procesados	630/3199	archivos...	(2255530 registros)
Procesados	640/3199	archivos...	(2259480 registros)
Procesados	650/3199	archivos...	(2292259 registros)
Procesados	660/3199	archivos...	(2410117 registros)
Procesados	670/3199	archivos...	(2513121 registros)
Procesados	680/3199	archivos...	(2547015 registros)
Procesados	690/3199	archivos...	(2564965 registros)
Procesados	700/3199	archivos...	(2657043 registros)
Procesados	710/3199	archivos...	(2746465 registros)
Procesados	720/3199	archivos...	(2781610 registros)
Procesados	730/3199	archivos...	(2794346 registros)
Procesados	740/3199	archivos...	(2892573 registros)
Procesados	750/3199	archivos...	(3019349 registros)
Procesados	760/3199	archivos...	(3091206 registros)
Procesados	770/3199	archivos...	(3102579 registros)
Procesados	780/3199	archivos...	(3177498 registros)
Procesados	790/3199	archivos...	(3289351 registros)
Procesados	800/3199	archivos...	(3359223 registros)
Procesados	810/3199	archivos...	(3372273 registros)
Procesados	820/3199	archivos...	(3386544 registros)
Procesados	830/3199	archivos...	(3399567 registros)
Procesados	840/3199	archivos...	(3410458 registros)
Procesados	850/3199	archivos...	(3423679 registros)
Procesados	860/3199	archivos...	(3434415 registros)
Procesados	870/3199	archivos...	(3519963 registros)
Procesados	880/3199	archivos...	(3650595 registros)
Procesados	890/3199	archivos...	(3658132 registros)
Procesados	900/3199	archivos...	(3668836 registros)
Procesados	910/3199	archivos...	(3689526 registros)
Procesados	920/3199	archivos...	(3722619 registros)
Procesados	930/3199	archivos...	(3736340 registros)
Procesados	940/3199	archivos...	(3771634 registros)
Procesados	950/3199	archivos...	(3901667 registros)
Procesados	960/3199	archivos...	(4016598 registros)
Procesados	970/3199	archivos...	(4050096 registros)
Procesados	980/3199	archivos...	(4054776 registros)
Procesados	990/3199	archivos...	(4061202 registros)
Procesados	1000/3199	archivos...	(4067638 registros)
Procesados	1010/3199	archivos...	(4079334 registros)
Procesados	1020/3199	archivos...	(4136334 registros)
Procesados	1030/3199	archivos...	(4192258 registros)
Procesados	1040/3199	archivos...	(4212868 registros)
Procesados	1050/3199	archivos...	(4217813 registros)
Procesados	1060/3199	archivos...	(4226181 registros)
Procesados	1070/3199	archivos...	(4231290 registros)
Procesados	1080/3199	archivos...	(4237070 registros)
Procesados	1090/3199	archivos...	(4242972 registros)
Procesados	1100/3199	archivos...	(4261440 registros)
Procesados	1110/3199	archivos...	(4382267 registros)
Procesados	1120/3199	archivos...	(4510491 registros)
Procesados	1130/3199	archivos...	(4564956 registros)
Procesados	1140/3199	archivos...	(4574902 registros)
Procesados	1150/3199	archivos...	(4684663 registros)
Procesados	1160/3199	archivos...	(4829721 registros)
Procesados	1170/3199	archivos...	(4908889 registros)
Procesados	1180/3199	archivos...	(4916011 registros)

Procesados	1190/3199	archivos...	(4999027 registros)
Procesados	1200/3199	archivos...	(5139053 registros)
Procesados	1210/3199	archivos...	(5231426 registros)
Procesados	1220/3199	archivos...	(5243117 registros)
Procesados	1230/3199	archivos...	(5306924 registros)
Procesados	1240/3199	archivos...	(5451857 registros)
Procesados	1250/3199	archivos...	(5566308 registros)
Procesados	1260/3199	archivos...	(5590771 registros)
Procesados	1270/3199	archivos...	(5624639 registros)
Procesados	1280/3199	archivos...	(5728773 registros)
Procesados	1290/3199	archivos...	(5816970 registros)
Procesados	1300/3199	archivos...	(5847669 registros)
Procesados	1310/3199	archivos...	(5862679 registros)
Procesados	1320/3199	archivos...	(5880151 registros)
Procesados	1330/3199	archivos...	(5890724 registros)
Procesados	1340/3199	archivos...	(5901200 registros)
Procesados	1350/3199	archivos...	(5910245 registros)
Procesados	1360/3199	archivos...	(5946045 registros)
Procesados	1370/3199	archivos...	(6071671 registros)
Procesados	1380/3199	archivos...	(6186082 registros)
Procesados	1390/3199	archivos...	(6226324 registros)
Procesados	1400/3199	archivos...	(6249081 registros)
Procesados	1410/3199	archivos...	(6376836 registros)
Procesados	1420/3199	archivos...	(6511593 registros)
Procesados	1430/3199	archivos...	(6571519 registros)
Procesados	1440/3199	archivos...	(6586391 registros)
Procesados	1450/3199	archivos...	(6696154 registros)
Procesados	1460/3199	archivos...	(6838140 registros)
Procesados	1470/3199	archivos...	(6918006 registros)
Procesados	1480/3199	archivos...	(6930407 registros)
Procesados	1490/3199	archivos...	(7021310 registros)
Procesados	1500/3199	archivos...	(7167449 registros)
Procesados	1510/3199	archivos...	(7266521 registros)
Procesados	1520/3199	archivos...	(7282326 registros)
Procesados	1530/3199	archivos...	(7328736 registros)
Procesados	1540/3199	archivos...	(7410437 registros)
Procesados	1550/3199	archivos...	(7476180 registros)
Procesados	1560/3199	archivos...	(7498188 registros)
Procesados	1570/3199	archivos...	(7512827 registros)
Procesados	1580/3199	archivos...	(7528121 registros)
Procesados	1590/3199	archivos...	(7540506 registros)
Procesados	1600/3199	archivos...	(7556707 registros)
Procesados	1610/3199	archivos...	(7568957 registros)
Procesados	1620/3199	archivos...	(7639849 registros)
Procesados	1630/3199	archivos...	(7776120 registros)
Procesados	1640/3199	archivos...	(7879980 registros)
Procesados	1650/3199	archivos...	(7902264 registros)
Procesados	1660/3199	archivos...	(7957205 registros)
Procesados	1670/3199	archivos...	(8099744 registros)
Procesados	1680/3199	archivos...	(8220614 registros)
Procesados	1690/3199	archivos...	(8256885 registros)
Procesados	1700/3199	archivos...	(8292559 registros)
Procesados	1710/3199	archivos...	(8423793 registros)
Procesados	1720/3199	archivos...	(8547220 registros)
Procesados	1730/3199	archivos...	(8600299 registros)
Procesados	1740/3199	archivos...	(8622579 registros)
Procesados	1750/3199	archivos...	(8740636 registros)
Procesados	1760/3199	archivos...	(8874263 registros)
Procesados	1770/3199	archivos...	(8943698 registros)
Procesados	1780/3199	archivos...	(8960209 registros)

Procesados	1790/3199	archivos...	(9046977 registros)
Procesados	1800/3199	archivos...	(9162746 registros)
Procesados	1810/3199	archivos...	(9230322 registros)
Procesados	1820/3199	archivos...	(9245818 registros)
Procesados	1830/3199	archivos...	(9259049 registros)
Procesados	1840/3199	archivos...	(9272704 registros)
Procesados	1850/3199	archivos...	(9284544 registros)
Procesados	1860/3199	archivos...	(9299353 registros)
Procesados	1870/3199	archivos...	(9311217 registros)
Procesados	1880/3199	archivos...	(9402445 registros)
Procesados	1890/3199	archivos...	(9528205 registros)
Procesados	1900/3199	archivos...	(9588422 registros)
Procesados	1910/3199	archivos...	(9604833 registros)
Procesados	1920/3199	archivos...	(9700272 registros)
Procesados	1930/3199	archivos...	(9826475 registros)
Procesados	1940/3199	archivos...	(9900473 registros)
Procesados	1950/3199	archivos...	(9913950 registros)
Procesados	1960/3199	archivos...	(9989058 registros)
Procesados	1970/3199	archivos...	(10114091 registros)
Procesados	1980/3199	archivos...	(10195761 registros)
Procesados	1990/3199	archivos...	(10210905 registros)
Procesados	2000/3199	archivos...	(10277607 registros)
Procesados	2010/3199	archivos...	(10401501 registros)
Procesados	2020/3199	archivos...	(10496111 registros)
Procesados	2030/3199	archivos...	(10518715 registros)
Procesados	2040/3199	archivos...	(10564542 registros)
Procesados	2050/3199	archivos...	(10667442 registros)
Procesados	2060/3199	archivos...	(10747071 registros)
Procesados	2070/3199	archivos...	(10772637 registros)
Procesados	2080/3199	archivos...	(10788201 registros)
Procesados	2090/3199	archivos...	(10807440 registros)
Procesados	2100/3199	archivos...	(10820851 registros)
Procesados	2110/3199	archivos...	(10841189 registros)
Procesados	2120/3199	archivos...	(10857243 registros)
Procesados	2130/3199	archivos...	(10898940 registros)
Procesados	2140/3199	archivos...	(10994326 registros)
Procesados	2150/3199	archivos...	(11079134 registros)
Procesados	2160/3199	archivos...	(11110483 registros)
Procesados	2170/3199	archivos...	(11140916 registros)
Procesados	2180/3199	archivos...	(11237129 registros)
Procesados	2190/3199	archivos...	(11298364 registros)
Procesados	2200/3199	archivos...	(11314817 registros)
Procesados	2210/3199	archivos...	(11387746 registros)
Procesados	2220/3199	archivos...	(11487196 registros)
Procesados	2230/3199	archivos...	(11558316 registros)
Procesados	2240/3199	archivos...	(11574939 registros)
Procesados	2250/3199	archivos...	(11632032 registros)
Procesados	2260/3199	archivos...	(11733682 registros)
Procesados	2270/3199	archivos...	(11815565 registros)
Procesados	2280/3199	archivos...	(11837959 registros)
Procesados	2290/3199	archivos...	(11879991 registros)
Procesados	2300/3199	archivos...	(11968766 registros)
Procesados	2310/3199	archivos...	(12042547 registros)
Procesados	2320/3199	archivos...	(12065885 registros)
Procesados	2330/3199	archivos...	(12081134 registros)
Procesados	2340/3199	archivos...	(12099135 registros)
Procesados	2350/3199	archivos...	(12113195 registros)
Procesados	2360/3199	archivos...	(12131201 registros)
Procesados	2370/3199	archivos...	(12146408 registros)
Procesados	2380/3199	archivos...	(12161949 registros)

Procesados	2390/3199	archivos...	(12186689	registros)
Procesados	2400/3199	archivos...	(12211565	registros)
Procesados	2410/3199	archivos...	(12230549	registros)
Procesados	2420/3199	archivos...	(12257134	registros)
Procesados	2430/3199	archivos...	(12336354	registros)
Procesados	2440/3199	archivos...	(12419117	registros)
Procesados	2450/3199	archivos...	(12456960	registros)
Procesados	2460/3199	archivos...	(12476920	registros)
Procesados	2470/3199	archivos...	(12536203	registros)
Procesados	2480/3199	archivos...	(12609413	registros)
Procesados	2490/3199	archivos...	(12655878	registros)
Procesados	2500/3199	archivos...	(12672031	registros)
Procesados	2510/3199	archivos...	(12723439	registros)
Procesados	2520/3199	archivos...	(12797041	registros)
Procesados	2530/3199	archivos...	(12851868	registros)
Procesados	2540/3199	archivos...	(12866144	registros)
Procesados	2550/3199	archivos...	(12912105	registros)
Procesados	2560/3199	archivos...	(12981218	registros)
Procesados	2570/3199	archivos...	(13033500	registros)
Procesados	2580/3199	archivos...	(13049230	registros)
Procesados	2590/3199	archivos...	(13067947	registros)
Procesados	2600/3199	archivos...	(13086562	registros)
Procesados	2610/3199	archivos...	(13101478	registros)
Procesados	2620/3199	archivos...	(13122697	registros)
Procesados	2630/3199	archivos...	(13141020	registros)
Procesados	2640/3199	archivos...	(13168116	registros)
Procesados	2650/3199	archivos...	(13231250	registros)
Procesados	2660/3199	archivos...	(13247186	registros)
Procesados	2670/3199	archivos...	(13306081	registros)
Procesados	2680/3199	archivos...	(13334586	registros)
Procesados	2690/3199	archivos...	(13374171	registros)
Procesados	2700/3199	archivos...	(13420772	registros)
Procesados	2710/3199	archivos...	(13437116	registros)
Procesados	2720/3199	archivos...	(13489063	registros)
Procesados	2730/3199	archivos...	(13505525	registros)
Procesados	2740/3199	archivos...	(13521342	registros)
Procesados	2750/3199	archivos...	(13537461	registros)
Procesados	2760/3199	archivos...	(13551355	registros)
Procesados	2770/3199	archivos...	(13568111	registros)
Procesados	2780/3199	archivos...	(13581855	registros)
Procesados	2790/3199	archivos...	(13616711	registros)
Procesados	2800/3199	archivos...	(13634537	registros)
Procesados	2810/3199	archivos...	(13649090	registros)
Procesados	2820/3199	archivos...	(13665930	registros)
Procesados	2830/3199	archivos...	(13683614	registros)
Procesados	2840/3199	archivos...	(13735176	registros)
Procesados	2850/3199	archivos...	(13751766	registros)
Procesados	2860/3199	archivos...	(13796117	registros)
Procesados	2870/3199	archivos...	(13820203	registros)
Procesados	2880/3199	archivos...	(13850156	registros)
Procesados	2890/3199	archivos...	(13881684	registros)
Procesados	2900/3199	archivos...	(13896800	registros)
Procesados	2910/3199	archivos...	(13914259	registros)
Procesados	2920/3199	archivos...	(13929297	registros)
Procesados	2930/3199	archivos...	(13944955	registros)
Procesados	2940/3199	archivos...	(13960832	registros)
Procesados	2950/3199	archivos...	(13991416	registros)
Procesados	2960/3199	archivos...	(14027794	registros)
Procesados	2970/3199	archivos...	(14044192	registros)
Procesados	2980/3199	archivos...	(14091176	registros)

```

Procesados 2990/3199 archivos... (14108884 registros)
Procesados 3000/3199 archivos... (14147718 registros)
Procesados 3010/3199 archivos... (14174720 registros)
Procesados 3020/3199 archivos... (14197186 registros)
Procesados 3030/3199 archivos... (14239524 registros)
Procesados 3040/3199 archivos... (14255254 registros)
Procesados 3050/3199 archivos... (14294720 registros)
Procesados 3060/3199 archivos... (14313829 registros)
Procesados 3070/3199 archivos... (14329251 registros)
Procesados 3080/3199 archivos... (14344947 registros)
Procesados 3090/3199 archivos... (14359608 registros)
Procesados 3100/3199 archivos... (14375142 registros)
Procesados 3110/3199 archivos... (14389646 registros)
Procesados 3120/3199 archivos... (14422660 registros)
Procesados 3130/3199 archivos... (14446649 registros)
Procesados 3140/3199 archivos... (14467770 registros)
Procesados 3150/3199 archivos... (14505695 registros)
Procesados 3160/3199 archivos... (14521629 registros)
Procesados 3170/3199 archivos... (14562051 registros)
Procesados 3180/3199 archivos... (14582921 registros)
Procesados 3190/3199 archivos... (14611443 registros)

```

✅ Cargados 14640925 registros de 3199 archivos

💾 Memoria: 26207.45 MB

=====

🌀 Clientes cargados: 14,640,925 registros

=====

🔍 3. Exploración Inicial

Veamos qué información tenemos disponible.

```

In [6]: # Ver primeras filas de APs
print("📁 MUESTRA DE ACCESS POINTS:")
df_aps.head()

```

📁 MUESTRA DE ACCESS POINTS:

Out [6]:

	ap_deployment_mode	ap_group	client_count	cluster_id	controller_name	cpu
--	--------------------	----------	--------------	------------	-----------------	-----

0	IAP	None	2			
---	-----	------	---	--	--	--

1	IAP	None	9			
---	-----	------	---	--	--	--

2	IAP	None	3			
---	-----	------	---	--	--	--

3	IAP	None	0			
---	-----	------	---	--	--	--

4	IAP	None	0			
---	-----	------	---	--	--	--

5 rows × 33 columns

```
In [17]: # Ver primeras filas de Clientes
print("📱 MUESTRA DE CLIENTES:")
df_clients.head()
```

📱 MUESTRA DE CLIENTES:

Out [17]:

	associated_device	associated_device_mac	associated_device_name	authentica
--	-------------------	-----------------------	------------------------	------------

0	AP_8e2d9933ec92	AP_5cdc80c05afc	AP-CEDU26	MAC Auth
---	-----------------	-----------------	-----------	----------

1	AP_f6e291deb8b9	AP_b509bd7bc558	AP-VET49	MAC Auth
---	-----------------	-----------------	----------	----------

2	AP_875979a461af	AP_e8c7358799e2	AP-LLET39	MAC Auth
---	-----------------	-----------------	-----------	----------

3	AP_332aab2f3da0	AP_35f1fd725406	AP-SAB-S239	MAC Auth
---	-----------------	-----------------	-------------	----------

4	AP_332aab2f3da0	AP_35f1fd725406	AP-SAB-S239	MAC Auth
---	-----------------	-----------------	-------------	----------

5 rows × 46 columns

```
In [7]: # Información del dataset de APs
print_dataset_summary(df_aps, "Access Points")
```

RESUMEN: Access Points

Registros totales: 2,663,195
 Columnas: 33
 Memoria: 3572.42 MB

Rango temporal:
 Inicio: 2024-10-23 16:52:06
 Fin: 2025-04-29 12:56:59
 Días: 187

Columnas disponibles:

- ap_deployment_mode (100.0% completo)
- ap_group (26.4% completo)
- client_count (100.0% completo)
- cluster_id (100.0% completo)
- controller_name (100.0% completo)
- cpu_utilization (100.0% completo)
- firmware_version (100.0% completo)
- gateway_cluster_id (100.0% completo)
- gateway_cluster_name (100.0% completo)
- group_name (100.0% completo)
- ... y 23 columnas más

```
In [8]: # Información del dataset de Clientes
print_dataset_summary(df_clients, "Clientes")
```

RESUMEN: Clientes

Registros totales: 14,640,925
 Columnas: 46
 Memoria: 27794.72 MB

Rango temporal:
 Inicio: 2025-03-28 12:08:19.808000
 Fin: 2025-07-10 18:59:50.986000
 Días: 104

Columnas disponibles:

- associated_device (100.0% completo)
- associated_device_mac (100.0% completo)
- associated_device_name (100.0% completo)
- authentication_type (97.9% completo)
- band (96.1% completo)
- channel (96.1% completo)
- client_category (97.3% completo)
- client_type (100.0% completo)
- connected_device_type (100.0% completo)
- connection (94.8% completo)
- ... y 36 columnas más

4. Estadísticas Básicas

4.1 Access Points

```
In [10]: # Número de APs únicos
unique_aps = df_aps['name'].nunique()
print(f"📶 APs únicos en el campus: {unique_aps}")

# Estado de los APs
print("\n🔌 Estado de los APs:")
print(df_aps['status'].value_counts())

# APs más populares (por número de clientes)
print("\n👤 Top 10 APs por número de clientes conectados:")
top_aps = df_aps.nlargest(10, 'client_count')[['name', 'client_count', 's
print(top_aps)
```

📶 APs únicos en el campus: 1147

🔌 Estado de los APs:

```
status
Up      2513363
Down    149832
Name: count, dtype: int64
```

👤 Top 10 APs por número de clientes conectados:

	name	client_count	status
238144	AP-BIBCIE01	159	Up
1916029	AP-LLET01	159	Up
237002	AP-BIBCIE01	158	Up
239286	AP-BIBCIE01	158	Up
242712	AP-BIBCIE01	158	Up
230150	AP-BIBCIE01	156	Up
241570	AP-BIBCIE01	156	Up
231292	AP-BIBCIE01	154	Up
92360	AP-BIBCIE01	153	Up
234718	AP-BIBCIE01	153	Up

4.2 Clientes/Dispositivos

```
In [11]: # Dispositivos únicos
unique_devices = df_clients['macaddr'].nunique()
print(f"📱 Dispositivos únicos: {unique_devices:,}")

# Tipos de red
print("\n🌐 Distribución por tipo de red:")
print(df_clients['network'].value_counts())

# Sistemas operativos
print("\n🖥️ Top 10 sistemas operativos:")
print(df_clients['os_type'].value_counts().head(10))

# Categorías de dispositivos
print("\n📦 Categorías de dispositivos:")
print(df_clients['client_category'].value_counts())
```


📱 Dispositivos únicos: 202,340

🌐 Distribución por tipo de red:

network	
UAB	10803753
eduroam	3700810
IdC	105572
IdC5GHz	11386
eduroamt	160

Name: count, dtype: int64

🖥️ Top 10 sistemas operativos:

os_type	
Apple iOS Device	4454656
Windows 8/10/11	2974238
Android	2911933
Mac OS	1759330
Apple iPhone	409294
Samsung Android	355447
--	295454
Xiaomi Device	277448
Windows	201706
Apple iPad	98738

Name: count, dtype: int64

📦 Categorías de dispositivos:

client_category	
SmartDevice	8653083
Computer	5096354
Home Audio/Video Equipment	189302
Printer	113898
Router	62655
IoT	57883
Generic	49002
Embedded	9211
Game Console	6522
Lighting	4239
Monitoring Devices	3769
Medical Device	388
Audio/Video Devices	97
SetTop Box	3
Point of Sale devices	1

Name: count, dtype: int64

📈 5. Visualizaciones Básicas

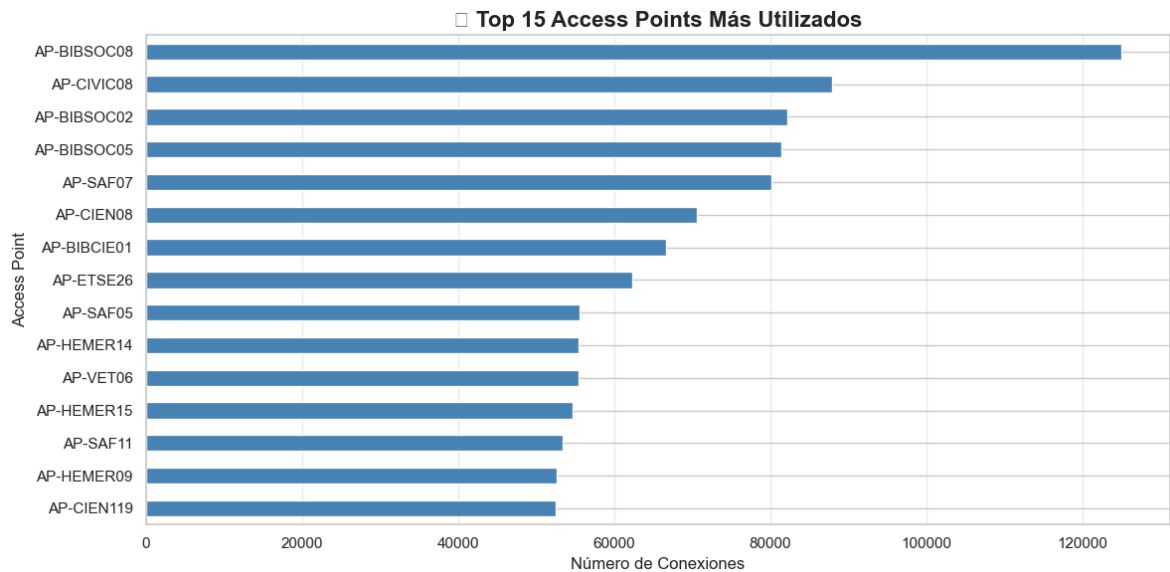
5.1 Zonas "Hotspot" - APs más utilizados

```
In [12]: # Top 15 APs más utilizados
top_aps_usage = df_clients['associated_device_name'].value_counts().head(15)

plt.figure(figsize=(12, 6))
top_aps_usage.plot(kind='barh', color='steelblue')
plt.title('📍 Top 15 Access Points Más Utilizados', fontsize=16, fontweight='bold')
plt.xlabel('Número de Conexiones', fontsize=12)
plt.ylabel('Access Point', fontsize=12)
plt.gca().invert_yaxis()
plt.grid(axis='x', alpha=0.3)
```

```
plt.tight_layout()
plt.show()
```

```
print(f"\n🔥 El AP más usado es: {top_aps_usage.index[0]} con {top_aps_us
```



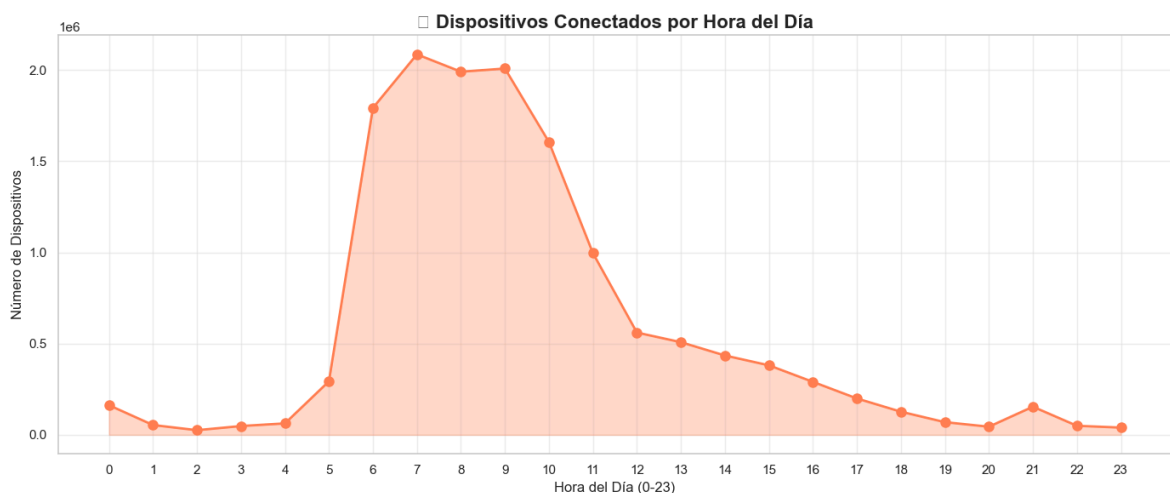
🔥 El AP más usado es: AP-BIBSOC08 con 124,985 conexiones

5.2 Patrones Temporales - Actividad por Hora

```
In [13]: # Actividad por hora del día
hourly_activity = df_clients.groupby('hour').size()

plt.figure(figsize=(14, 6))
plt.plot(hourly_activity.index, hourly_activity.values, marker='o', linew
plt.fill_between(hourly_activity.index, hourly_activity.values, alpha=0.3
plt.title('🕒 Dispositivos Conectados por Hora del Día', fontsize=16, fon
plt.xlabel('Hora del Día (0-23)', fontsize=12)
plt.ylabel('Número de Dispositivos', fontsize=12)
plt.xticks(range(0, 24))
plt.grid(alpha=0.3)
plt.tight_layout()
plt.show()

# Identificar hora pico
peak_hour = hourly_activity.idxmax()
peak_devices = hourly_activity.max()
print(f"\n🕒 Hora pico: {peak_hour}:00 con {peak_devices:,} dispositivos
```



🕒 Hora pico: 7.0:00 con 2,086,370 dispositivos conectados

5.3 Distribución de Calidad de Señal

```
In [14]: # Distribución de signal_strength (1-5)
fig, axes = plt.subplots(1, 2, figsize=(14, 5))

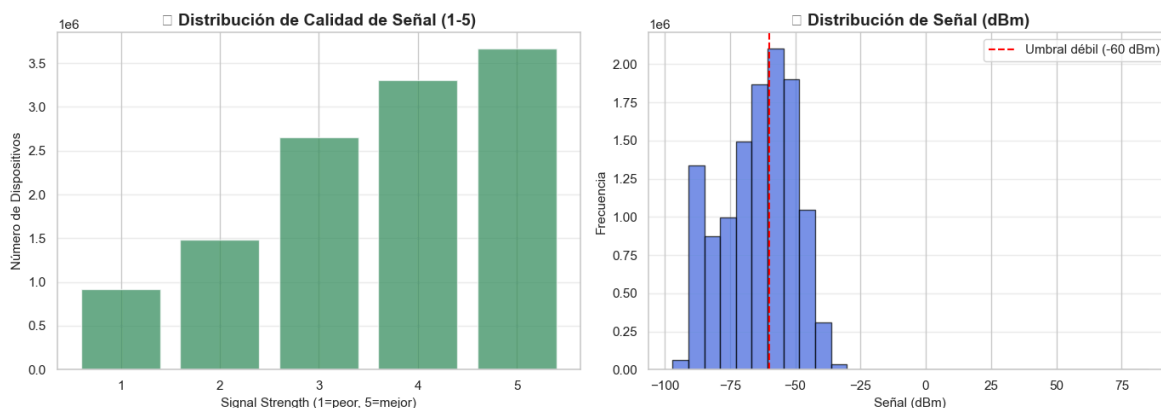
# Gráfico 1: Signal Strength (1-5)
signal_counts = df_clients['signal_strength'].value_counts().sort_index()
axes[0].bar(signal_counts.index, signal_counts.values, color='seagreen',
axes[0].set_title('📊 Distribución de Calidad de Señal (1-5)', fontsize=14)
axes[0].set_xlabel('Signal Strength (1=peor, 5=mejor)', fontsize=11)
axes[0].set_ylabel('Número de Dispositivos', fontsize=11)
axes[0].set_xticks([1, 2, 3, 4, 5])
axes[0].grid(axis='y', alpha=0.3)

# Gráfico 2: Signal dB (histograma)
axes[1].hist(df_clients['signal_db'].dropna(), bins=30, color='royalblue')
axes[1].set_title('📊 Distribución de Señal (dBm)', fontsize=14, fontweig
axes[1].set_xlabel('Señal (dBm)', fontsize=11)
axes[1].set_ylabel('Frecuencia', fontsize=11)
axes[1].axvline(x=-60, color='red', linestyle='--', label='Umbral débil (
axes[1].legend()
axes[1].grid(axis='y', alpha=0.3)

plt.tight_layout()
plt.show()

# Estadísticas
avg_signal = df_clients['signal_db'].mean()
weak_signal = (df_clients['signal_db'] < -60).sum()
pct_weak = (weak_signal / len(df_clients)) * 100

print(f"\n📶 Señal promedio: {avg_signal:.1f} dBm")
print(f"⚠️ Dispositivos con señal débil (<-60 dBm): {weak_signal:,} ({pct
```



📶 Señal promedio: -64.1 dBm

⚠️ Dispositivos con señal débil (<-60 dBm): 6,629,869 (45.3%)

5.4 Tipos de Red (UAB vs eduroam)

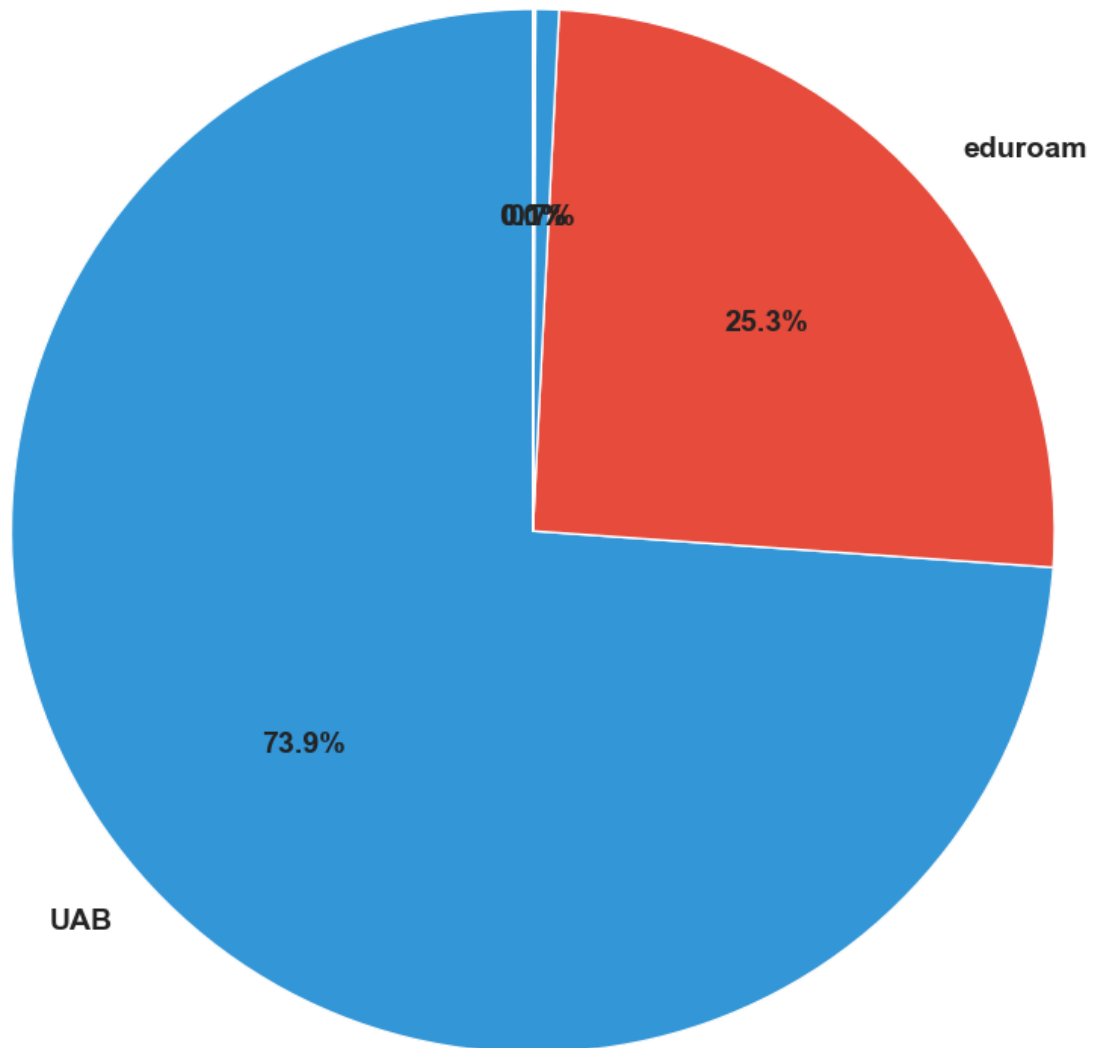
```
In [15]: # Distribución por tipo de red
network_counts = df_clients['network'].value_counts()

plt.figure(figsize=(8, 8))
colors = ['#3498db', '#e74c3c']
```

```
plt.pie(network_counts.values, labels=network_counts.index, autopct='%1.1f',
        startangle=90, colors=colors, textprops={'fontsize': 14, 'weight': 'bold'})
plt.title('🌐 Distribución por Tipo de Red', fontsize=16, fontweight='bold')
plt.axis('equal')
plt.tight_layout()
plt.show()

print("\n📊 Desglose por red:")
for network, count in network_counts.items():
    pct = (count / len(df_clients)) * 100
    print(f"    {network}: {count:,} dispositivos ({pct:.1f}%)")
```

□ Distribución por Tipo de Red



📊 Desglose por red:

- UAB: 10,803,753 dispositivos (73.8%)
- eduroam: 3,700,810 dispositivos (25.3%)
- IdC: 105,572 dispositivos (0.7%)
- IdC5GHz: 11,386 dispositivos (0.1%)
- eduroamt: 160 dispositivos (0.0%)

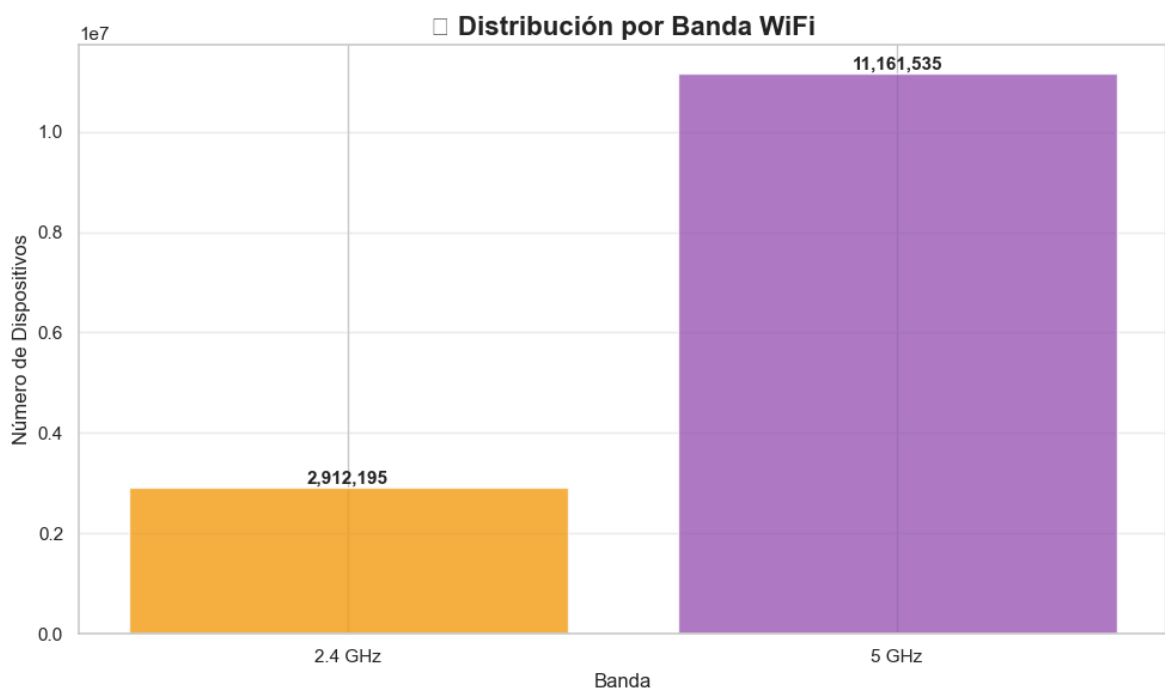
5.5 Bandas WiFi (2.4 GHz vs 5 GHz)

```
In [16]: # Distribución por banda
band_counts = df_clients['band'].value_counts().sort_index()
band_labels = {2.4: '2.4 GHz', 5: '5 GHz', 6: '6 GHz'}
```

```
plt.figure(figsize=(10, 6))
colors_band = ['#f39c12', '#9b59b6', '#1abc9c']
bars = plt.bar([band_labels.get(x, str(x)) for x in band_counts.index],
               band_counts.values, color=colors_band[:len(band_counts)]),
plt.title('📶 Distribución por Banda WiFi', fontsize=16, fontweight='bold')
plt.xlabel('Banda', fontsize=12)
plt.ylabel('Número de Dispositivos', fontsize=12)
plt.grid(axis='y', alpha=0.3)

# Añadir valores encima de las barras
for bar in bars:
    height = bar.get_height()
    plt.text(bar.get_x() + bar.get_width()/2., height,
             f'{int(height):,}', ha='center', va='bottom', fontsize=11, f

plt.tight_layout()
plt.show()
```



🎯 6. Insights y Conclusiones

Basándote en las visualizaciones anteriores, responde:

1. ¿Cuál es el AP más utilizado? ¿En qué edificio está?

Tu respuesta aquí: AP-BIBS0C08, en el edificio de la biblioteca de ciencias sociales.

2. ¿A qué hora hay más actividad? ¿Por qué crees que es?

Tu respuesta aquí: A las 7 am

3. ¿Qué porcentaje de dispositivos tiene señal débil?

Tu respuesta aquí: 45.3%

4. ¿Cuál es la red más utilizada (UAB o eduroam)?

Tu respuesta aquí: UAB es mucho más utilizada que eduroam

5. ¿Qué banda WiFi es más popular (2.4 o 5 GHz)? ¿Por qué?

Tu respuesta aquí: 5Ghz, se utiliza para evitar interferencias y saturación.

7. Próximos Pasos

¡Felicidades! Has completado el análisis básico. Ahora puedes:

Ideas para Expandir tu Análisis:

1. **Cargar más datos:** Cambia `max_files=100` o `max_files=None` para todo el dataset
2. **Análisis por edificio:**
 - Extraer el código del edificio del nombre del AP (ej: AP-VET71 → VET)
 - Comparar actividad entre edificios
3. **Análisis temporal más profundo:**
 - Comparar días laborables vs fines de semana
 - Identificar patrones semanales
4. **Problemas de conectividad:**
 - Identificar APs con peor señal
 - Analizar zonas problemáticas
5. **Dashboard interactivo:**
 - Usar Plotly para gráficos interactivos
 - Crear filtros por fecha/hora/edificio

¿Listo para el siguiente nivel?

👉 Continúa con: `02_intermediate_mobility.ipynb`

🚩 UAB THE HACK! 2025 - ¡Buena suerte! 🚀